PhytoKeys

PhytoKeys 241: 27–48 (2024)
DOI: 10.3897/phytokeys.241.118903

Research Article

Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny

Elif Yılmaz[1], David G. Mann[2], Romain Gastineau[1], Rosa Trobajo[3], Cüneyt Nadir Solak[4], Ewa Górecka[1], Monique Turmel[5], Claude Lemieux[5], Nesil Ertorun[6], Andrzej Witkowski[1*]

1 *Institute of Marine and Environmental Sciences, University of Szczecin, Mickiewicza 16A, PL70–383 Poland*
2 *Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, Scotland, UK*
3 *Marine and Continental Waters, Institute for Food and Agricultural Research and Technology (IRTA), Crta de Poble Nou Km 5.5, E-43540 La Ràpita, Catalunya, Spain*
4 *Department of Biology, Faculty of Science & Art, Dumlupınar University, 43000 Kütahya, Türkiye*
5 *Département de biochimie, de microbiologie et de bio-Informatique, Institut de Biologie Intégrative et des Systèmes, Université Laval, Québec, QC, Canada*
6 *Department of Biology, Science Faculty, Eskişehir Technical University, 26000 Eskişehir, Türkiye*

Corresponding author: Elif Yılmaz (elfyilmaz38@gmail.com)

Abstract

The current article describes *Navicula vanseea* **sp. nov.**, a new species of diatom from Lake Van, a highly alkaline lake in Eastern Anatolia (Türkiye). The description is based on light and scanning electron microscopy performed on two monoclonal cultures. The complete nuclear rRNA clusters and plastid genomes have been sequenced for these two strains and the complete mitogenome for one of them. The plastome of both strains shows the probable loss of a functional *ycf35* gene. They also exhibit two IB4 group I introns in their *rrl*, each encoding for a putative LAGLIDADG homing endonuclease, with the first L1917 IB4 intron reported amongst diatoms. The Maximum Likelihood phylogeny inferred from a concatenated alignment of *18S*, *rbcL* and *psbC* distinguishes *N. vanseea* sp. nov. from the morphologically similar species *Navicula cincta* and *Navicula microdigitoradiata*.

Key words: Group I intron, LAGLIDADG, mitogenome, Naviculaceae, plastome, pseudogene, soda lake

OPEN ACCESS

Academic editor: Bing Liu
Received: 16 January 2024
Accepted: 19 February 2024
Published: 8 April 2024

Citation: Yilmaz E, Mann DG, Gastineau R, Trobajo R, Solak CN, Górecka E, Turmel M, Lemieux C, Ertorun N, Witkowski A (2024) Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. PhytoKeys 241: 27–48. https://doi.org/10.3897/phytokeys.241.118903

Copyright: © Elif Yılmaz et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

* Deceased author.

Introduction

Lake Van is located in Eastern Anatolia, Turkey (Republic of Türkiye). It is Turkey's largest inland water body and also world's largest soda lake. The lake is surrounded by dormant volcanoes and its formation was a consequence of the eruption of the Nemrut stratovolcano (not to be confused with the Nemrut Mountain, also in Turkey), which is 2247 m above sea level. As a result of the erosion of volcanic rocks in the catchment and evaporation, the lake water is salty (21.4‰) and alkaline (155 m mEq−1, pH 9.81) (Glombitza et al. 2013; Ersoy Omeroglu et al. 2021). The lake is notable for its unusual chemistry,

which results from the constant losses of calcium as carbonate and of magnesium in the form of mineral phases rich in Mg–silica. Thus, the Mg cycle is closely related to the silica cycle, which is itself dependent on the production of biosilica by diatoms, eventually followed by the dissolution of their frustules (Reimer et al. 2009).

The genus *Navicula* is amongst the most species-rich genera of Bacillariophyceae, although this is partly because it was used for a long time as a 'catch-all' for simply structured, bilaterally symmetrical raphid diatoms. It was erected as early as 1822 by Bory in his 'Dictionnaire Classique d'Histoire Naturelle' (Bory de Saint-Vincent 1822). The name chosen by Bory refers to the shape of the cells, similar to the shuttle that was used for weaving. The cells are generally solitary and motile, although some species live in mucilage tubes (Millie and Wee 1981). Cells have two parietal chloroplasts. Their valves are symmetrical both apically and transapically and have rounded, acute or capitate ends. The central area is often distinctly expanded (Patrick 1959; Cox 1979; Round et al. 1990).

The only account ever published on the diatoms from Lake Van was written by Legler and Krasske (1940). Amongst the 24 species they recorded were three taxa of *Navicula*, all of them considered as varieties of *Navicula cryptocephala*, namely *Navicula cryptocephala* Kützing, 1844, *Navicula cryptocephala* var. *intermedia* Grunow 1880 and a taxon noted as *Navicula cryptocephala* var. *veneta* (Kützing) Grunow, which possibly corresponds to *Navicula cryptocephala* var. *veneta* (Kützing) Rabenhorst, 1864. Out of these three taxa, only *N. cryptocephala* Kützing, 1844 is still deemed to be valid. *Navicula cryptocephala* var. *intermedia* is considered a synonym of *Navicula capitatoradiata* H. Germain ex Gasse 1986 and *Navicula cryptocephala* var. *veneta* is now treated as an independent species, *Navicula veneta* Kützing 1844.

Navicula species are rather well documented in inland waters where they are known for their bioindicator potential (Lange-Bertalot 2001). For instance, *Navicula tripunctata* (O.F. Müller) Bory is a good indicator of eutrophic waters with an average to high electrolyte content, while taxa such as *Navicula gregaria* Donkin, 1861, *Navicula meulemansii* A.Mertens, A.Witkowski & Lange-Bertalot 2013 and *N. veneta* are common in brackish to electrolyte rich waters (Cox 1995; Adrienne Mertens et al. 2014).

Preliminary results from a new sampling campaign conducted in 2021 in Lake Van strongly suggested that the biodiversity of diatoms had been underestimated in the previous work of Legler and Krasske (1940). One illustration is a previously undocumented *Nitzschia*, *N. anatoliensis* Górecka, Gastineau & Solak (Solak et al. 2021), which would have probably been overlooked if it had not been for the combined use of microscopic and molecular tools. Amongst several other monoclonal cultures from the 2021 campaign, two contained strains of a *Navicula* species were identified, which is the subject of the present article. Although both strains were noticeably different in size, they were quickly proven to belong to the same new species.

The aim of the following article is to formally describe *Navicula vanseea* sp. nov. from Lake Van. The description combines the use of light microscopy (live specimen and cleaned frustules) and scanning electron microscopy. The complete cluster of nuclear ribosomal RNA genes and the complete plastid genome were obtained for both strains by means of next generation sequencing and

also the mitogenome of one of these strains. These results were included in a multigene Maximum Likelihood phylogeny which unambiguously separated *Navicula vanseea* sp. nov. from morphologically similar known species, whose differences with *Navicula vanseea* sp. nov. are discussed. As it was the first time that a L1917 group I intron with its putative LAGLIDADG homing endonuclease gene had been discovered in the plastid genome of a diatom, special attention has been paid to this feature, with a phylogeny of the putative LAGLIDADG protein being performed.

Materials and methods

Sampling, isolation and cultivation

Epilithic samples were collected by brushing rocks in the littoral of Lake Van in July 2021, in the vicinity of Erciş Municipality (Fig. 1). Samples were re-suspended in surface water from the lake in 50 ml tubes before being brought to the University of Szczecin for subsequent analyses. Samples were then transferred into Petri dishes containing sterile f/2 medium (Guillard 1975) modified to 18‰ salinity. Single cells were isolated by micropipette under an inverted Nikon Eclipse light microscope. Successive re-isolations were performed (at least 3 times) before the culture was considered monoclonal. Strains were later transferred into 250 ml Erlenmeyer flasks with modified f/2 medium. Cultures were maintained in active growth under a light intensity of 60 µmol photons m^{-2} s^{-1} and a photoperiod of 14 h light/10 h darkness). Two morphologically distinct clones with different cell sizes were registered in the Szczecin Culture Collection as SZCZEY2172 and SZCZEY2262.

Light and scanning electron microscopy

Pictures of living diatoms were taken using a Light Microscope (LM) Zeiss Axio Scope A1 (Carl Zeiss, Jena, Germany at a magnification of 1000× by transferring diatom cultures directly on to the glass slide.

To prepare cleaned frustules for microscopy, 5 ml of monoclonal cultures were transferred into 20 ml beakers with 10 ml of 10% hydrochloric acid (HCl). After 24h, samples were washed four times with distilled water then re-suspended in 30% hydrogen peroxide (H_2O_2) and boiled for about four hours. Finally, samples were washed again four times with distilled water. For LM, cleaned material was then air-dried on cover glasses and mounted on glass slide with Naphrax® (Brunel Microscopes Ltd., Chippenham, UK) solution and pictures were taken with the Zeiss Axio Scope A1. For SEM, a drop of cleaned sample was deposited on a Nuclepore Track-Etch membrane from Whatman (Maidstone, England). The membranes were air-dried overnight, mounted on aluminium stubs with carbon tape and coated with gold using a Q150T coater from Quorum Technologies (Laughton, U.K.). SEM observations were made at the Faculty of Chemical Technology and Engineering, Western Pomeranian University of Technology in Szczecin (Poland), using a Hitachi SU8020 (Tokyo, Japan) and Eskişehir Technical University (Türkiye) using a ZEISS Ultra microscope (Oberkochen, Germany).

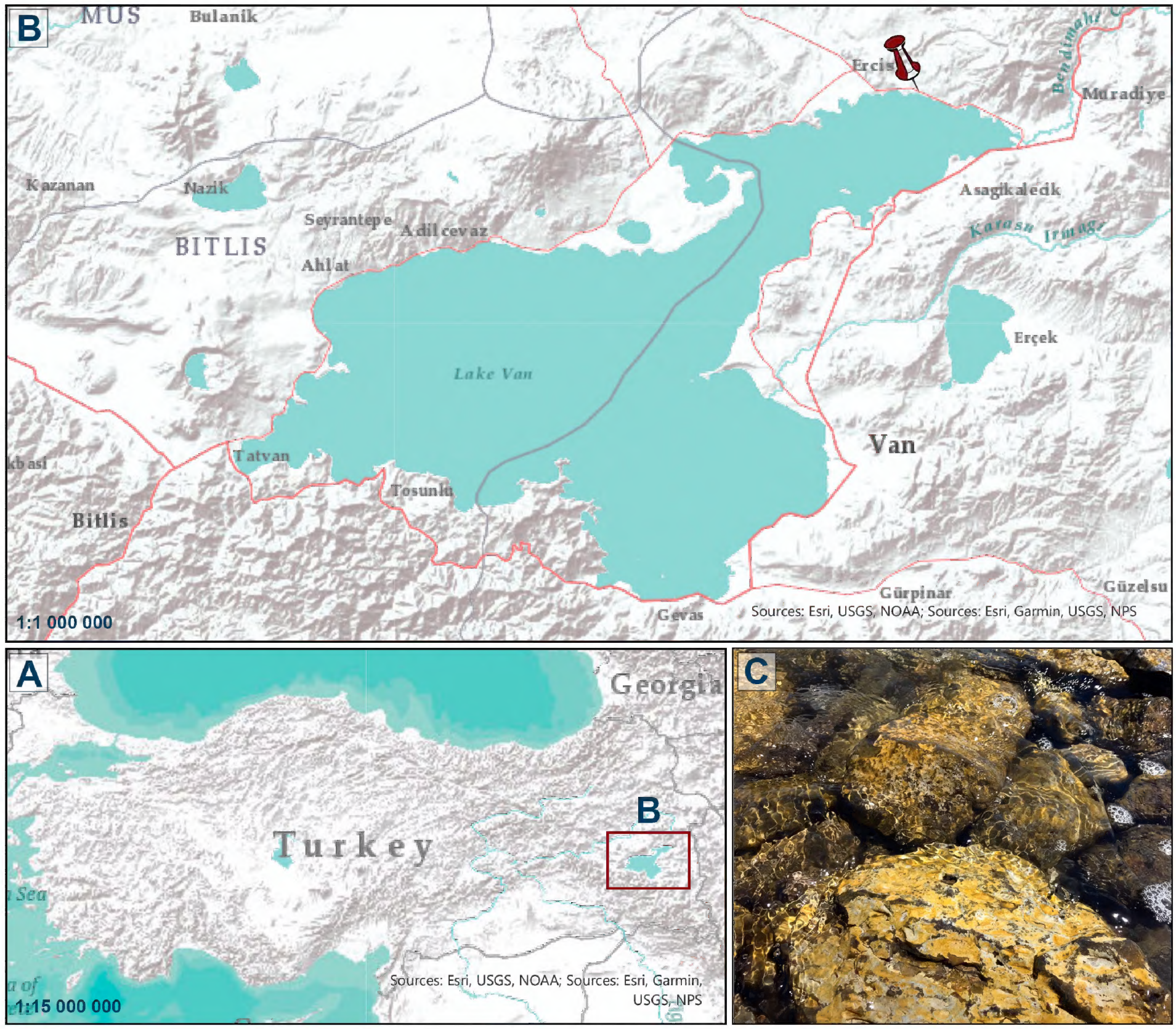

Figure 1. Map of the sampling location **A** location of Lake Van in Turkey. The red frame indicates the position of Lake Van **B** general view of the lake. The pin indicates the position of the sampling area **C** photo of the epilithic sampling area on the rock (Esri. (2023). ArcGIS Pro 3.1.0. Environmental Systems Research Institute).

Next generation sequencing and bioinformatic analyses

DNA was extracted from clones SZCZEY2172 and SZCZEY2262 using the protocol of Doyle and Doyle (1990). Total DNA was then sent to the Beijing Genomics Institute (BGI) in Shenzhen (China) to be sequenced on a DNBSEQ platform. For each clone, a total of ca. 40M clean 150 bp paired-end reads was obtained. Reads were assembled with a k-mer parameter of 125 using SPAdes 3.15.0 (Bankevich et al. 2012). Contigs of interest were retrieved by customised command-line BLASTn analyses as previously described (Dąbek et al. 2022; Gastineau et al. 2022). Consed (Gordon and Green 2013) was used to merge the different subunits of the plastome and when trying to circularise the mitogenome. Annotations were performed using the same tools as described in Gastineau et al. (2022). The maps of the organellar genomes were obtained from the OGDRAW online portal (Lohse et al. 2013). The different parts of the nuclear rRNA gene cluster were identified with the help of Rfam 14 (Kalvari et al. 2021).

Molecular phylogeny

The three gene datasets (*18S*, *rbcL* and *psbC*) already used in previous publications (Dąbek et al. 2017; Li et al. 2020; Górecka et al. 2021a;) were obtained and the corresponding genes from various Naviculaceae and *N. vanseea sp. nov*. appended. Sequences for *N. capitatoradiata*, *N. microdigitoradiata* Lange-Bertalot, 1993 and *N. cincta* (Ehrenberg) Ralfs 1861 were also added. However, it should be noted that these three species were represented in GenBank just by *18S* and *rbcL* (*N. capitatoradiata* and *N. cincta*) or *rbcL* only (*N. microdigitoradiata*). Genes were aligned separately with MAFFT 7 (Katoh and Standley 2013) with the -auto option and trimmed using trimAl (Capella-Gutiérrez et al. 2009) with the -automated1 option. The best model of evolution for each of these genes was selected with ModelTest-NG (Darriba et al. 2020) and were GTR+I+G4 (*psbC*), TIM3+I+G4 (*rbcL*) and TIM1+I+G4 (*18S*). Alignments were then concatenated using Phyutility 2.7.1 (Smith and Dunn 2008) for a final size of 3301 bp. A Maximum Likelihood phylogeny was constructed from the concatenated alignment using IQ-TREE 2.2.0 (Minh et al. 2020) with 1000 ultrafast bootstrap replicates and a dataset partitioned, based on the best models of evolution found for each gene. *Triparma pacifica* (Guillou & Chrétiennot-Dinet) Ichinomiya & Lopes dos Santos, 2016 was used as an outgroup.

For the phylogeny of the putative LAGLIDADG endonuclease proteins, protein sequences found in IB4 - L1917 and - L1931 introns presented in Lucas et al. (2001) and Haugen and Bhattacharya (2004) were downloaded from GenBank. Sequences from the current study, from Turmel et al. (2002), Haugen et al. (2007) and Lemieux et al. (2014) were appended. The IB2 – L1917 LAGLIDADG sequence from *Coxiella burnetii* (Derrick, 1939) Philip, 1948 was used as an outgroup (Raghavan et al. 2007). The phylogeny was conducted in a similar way to the multigene phylogeny, except that alignment was not trimmed. The best model of evolution returned by ModelTest-NG was LG+G4.

Results

Taxonomy

***Navicula vanseea* Yilmaz, Gastineau, Solak & Witkowski, sp. nov.**
Figs 2–4

Type material. *Holotype*: Slide number SZCZEY2172 in the collection of Andrzej Witkowski at the University of Szczecin, Poland. Valves representing the holotype population are illustrated in Fig. 2L.

***Isotype*:** Slide number TR_Erciş_Van_2021 deposited in Kütahya Dumlupınar University (Turkey).

Registration. http://phycobank.org/104542

Type locality. Erciş Van, Türkiye (38°59'47.3"N 43°24'15.3"E) collected by: Elif Yilmaz, 31 July 2021.

Etymology. The name given to the species refers to the German name of Lake Van (Vansee, the sea of Van) as it was used in the work of Legler and Krasske and is meant as a tribute to these authors and their work.

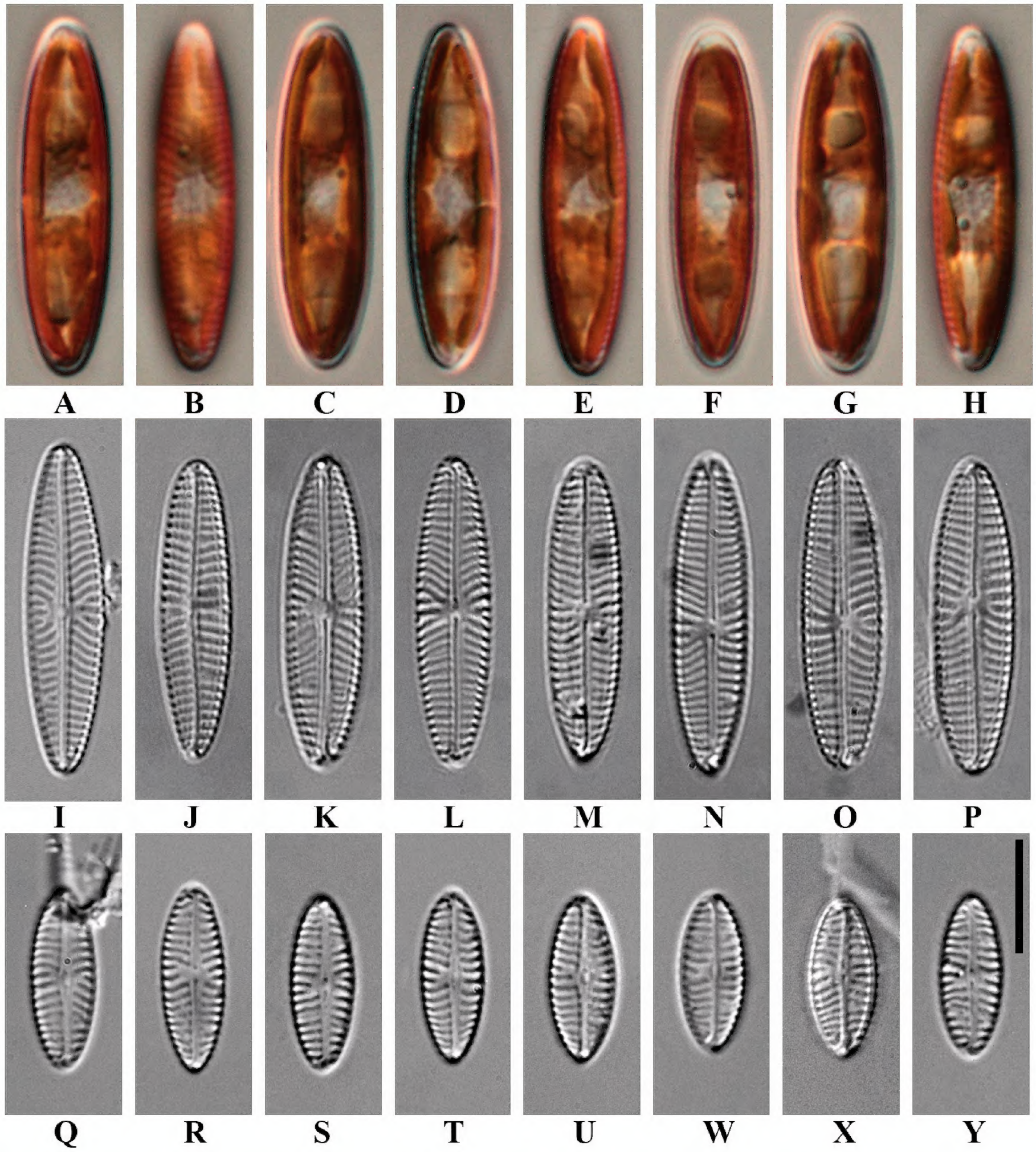

Figure 2. *Navicula vanseea* sp. nov. LM micrographs **A–H** *in vivo* pictures of *Navicula vanseea* sp. nov. SZCZEY2172 **I** LM image of a cleaned valve from wild material **J–P** cleaned valves of *Navicula vanseea* sp. nov. SZCZEY2172 **Q–Y** cleaned valves of *Navicula vanseea* sp. nov. SZCZEY2262 Scale bar: 10 µm.

Distribution and ecology. The taxon was exclusively observed within benthic epilithic assemblages in Lake Van (salinity 21.4‰ and pH 9.5).

Description. ***LM*** (Fig. 2A–Y) Valves: smaller specimens elliptic, tapering towards cuneately rounded apices, larger specimens linear-elliptic-lanceolate narrowly rounded, with narrowly rounded poles, which are occasionally slightly protracted (Fig. 2E, K, N). Valve dimensions (n = 39): length 11.0–28.0 µm, width 5.0–6.5 µm. Raphe filiform, straight. Central area small and rounded, axial area narrow. Striae strongly radiate, sometimes irregularly shortened around the central area, 12–13 in 10 µm, lineolae difficult to resolve in LM, ca. 50 in 10 µm.

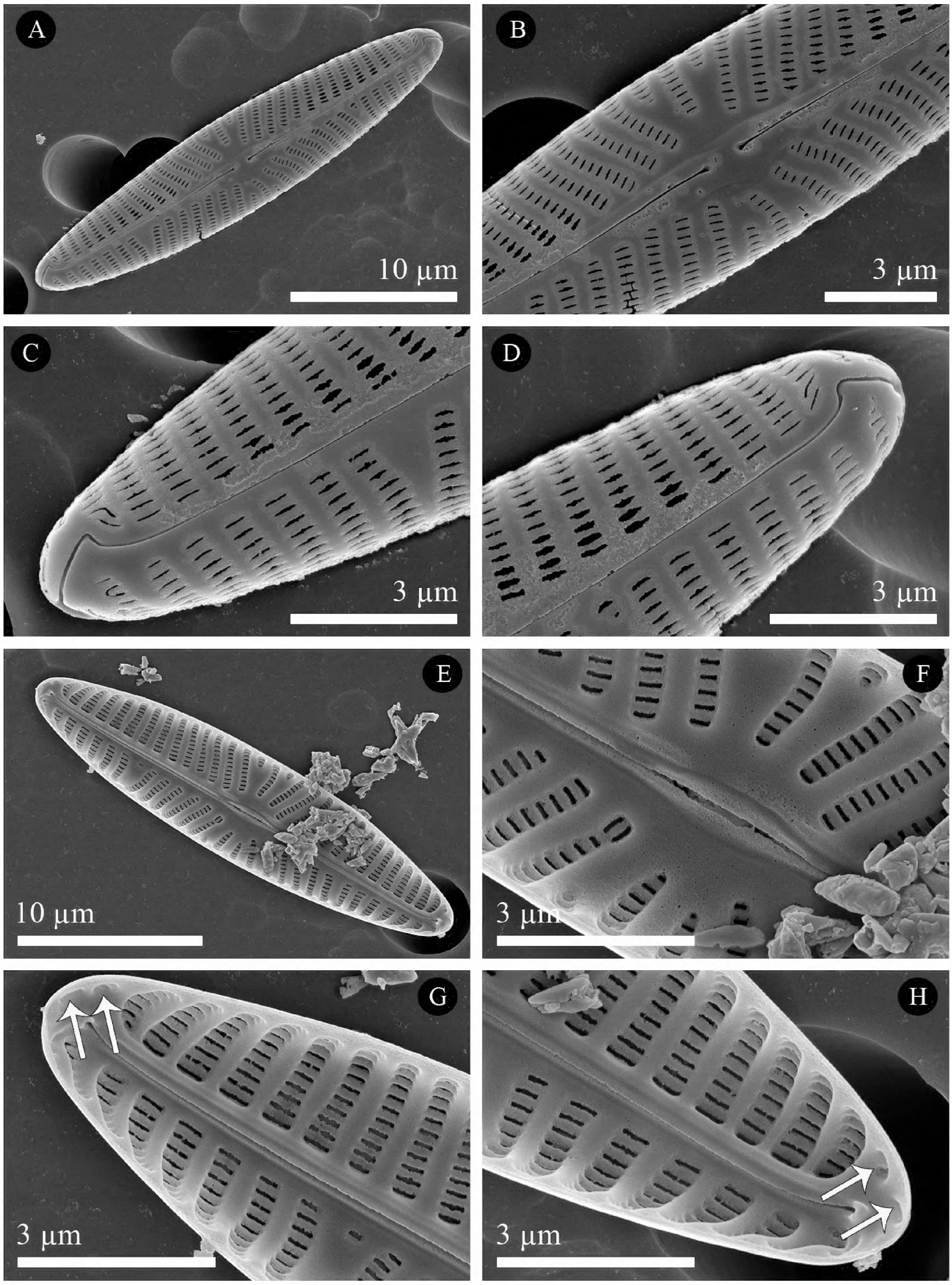

Figure 3. SEM micrographs of *Navicula vanseea* sp. nov. SZCZEY2172 **A** external view of the entire valve **B** details of central area showing simple, slightly drop-shaped proximal raphe endings and shortened striae **C, D** details of the two apices of a single valve showing the terminal fissures **E** internal view of the entire valve **F** details of central area showing filiform proximal raphe endings in a fusiform expansion of the raphe-sternum **G, H** details of apices showing well-developed helictoglossae showing two isolated lineolae (white arrows). Scale bars: 10 µm (**A, E**); 3 µm (**B–D, F–H**).

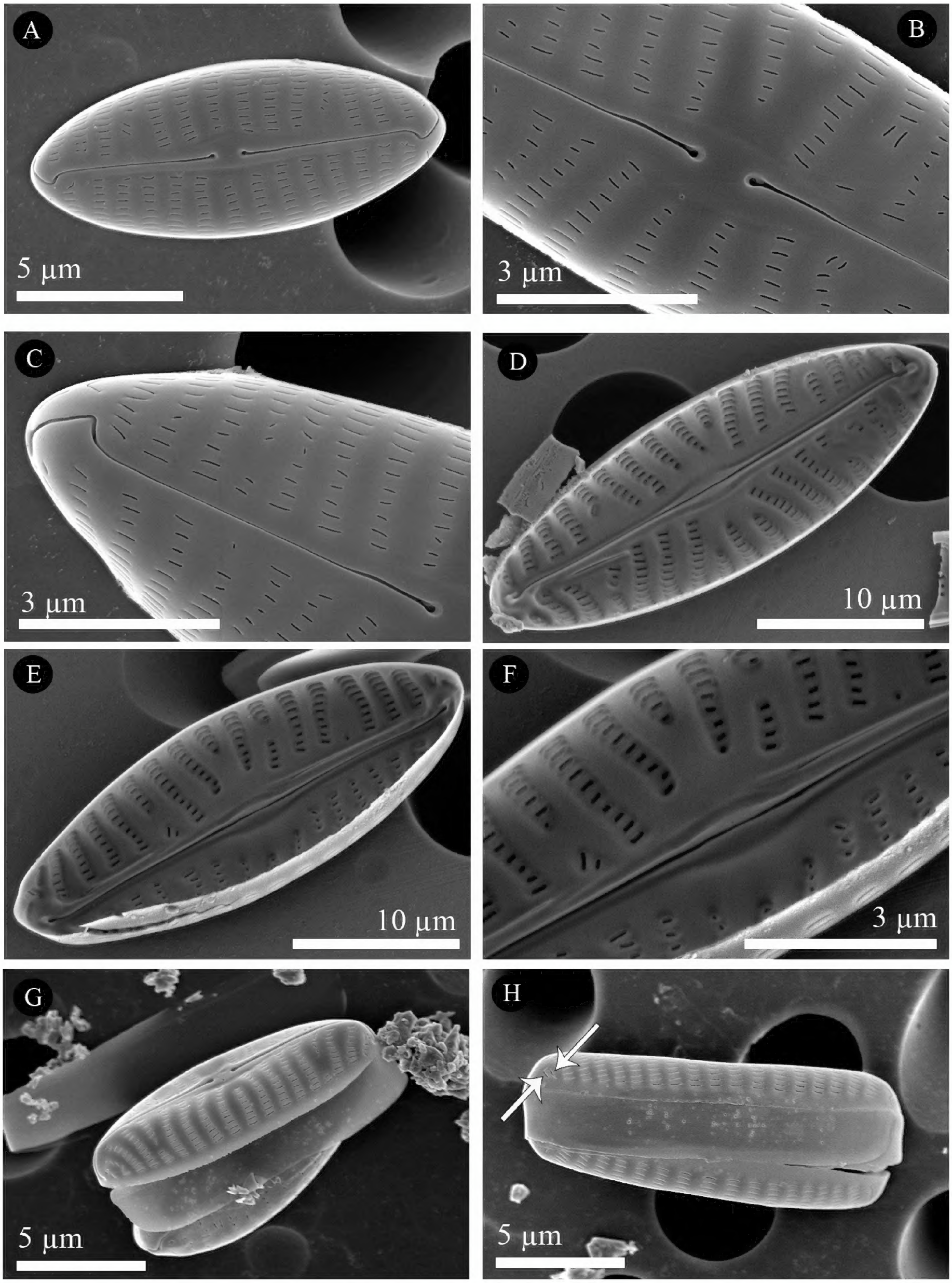

Figure 4. SEM micrographs of *Navicula vanseea* sp. nov. SZCZEY2262 **A** external view of the entire valve **B** details of central area showing simple proximal raphe endings and shortened striae **C** details of apex showing the terminal fissure **D, E** internal view of two entire valves, showing the central area and filiform proximal raphe endings **F** details of apex showing well-developed helictoglossae **G, H** girdle view of valves showing continuous areolation on mantle and two isolated lineolae (white arrows). Scale bars: 5 µm (**A, D, E, G, H**); 3 µm (**B, C, F**).

SEM External valve surface (Figs 3A–D, 4D–F): Valve surface flat (Fig. 3G, H), areolae apically elongated (Figs 3B–D, 4B). Raphe sternum slightly elevated above the valve face level (Fig. 4G). Axial area very narrow, central area very slightly expanded, small, asymmetric (Figs 3B, 4A, B). Proximal raphe endings drop-like, slightly deflected unilaterally (Fig. 4B). Distal raphe endings strongly hooked in the same direction (Fig. 3C, D, which are the two ends of the same valve and Fig. 4A).

SEM Internal valve surface (Figs 3E–H, 4D–F): valve surface slightly arched with transapical striae positioned in relatively deep grooves, bordered by virgae that become thicker towards to the centre of the valve (Figs 3H, 4D, F). Central area asymmetric, usually only slightly expanded (Figs 3E, F, 4F), but sometimes more strongly (Fig. 4D). The internal lineolae openings are slit-like (Fig. 3F–H), narrower than the vimines. Lineolae occluded by hymens (Fig. 4F); two isolated lineolae are present at the valve apex. Raphe sternum slightly widened at the centre to form a fusiform ridge enclosing the central raphe endings, which are simple, straight and separated (Figs 3F, 4F). Distally, the raphe terminates in well-developed helictoglossae (Figs 3G, H).

Genomics and phylogeny

The nuclear rRNA gene cluster

The complete rRNA gene cluster was sequenced for both clones and deposited in NCBI GenBank with accession numbers OR797294 (SZCZEY2172) and OR797293 (SZCZEY2262). The cluster is 4902 bp long, distributed as follows: 18S – 1792 bp, ITS1 – 195 bp, 5.8S – 155 bp, ITS2 – 260 bp, 28S – 2500 bp. Comparing the two clones, there was one single nucleotide polymorphism (SNP) found in the 18S (in the V2 region), three in the ITS1, one in the 5.8S, three in the ITS2 and two in the 28S (both in the D1/D2 region).

Mitochondrial genome

A 43997 bp contig corresponding to the mitochondrial genome was retrieved for strain SZCZEY2262, but could not be circularised because of the presence of repeated sequences at its ends. However, for easier reading, it is displayed as circular on the map (Fig. 5). The mitogenome encodes 34 protein-coding genes plus the conserved open-reading frame (ORF) *orf150* between *rps11* and *mttB* (Pogoda et al. 2019), two rRNA genes and 23 tRNA genes (GenBank: OR795084). The *nad11* gene is split into two distinct subunits, separated from each other by two protein-coding genes, two rRNA and one tRNA. In the repeated part of the genome, there are two copies of the same ORF, *orf145*. There is a 767-bp group I intron in the *rnl* gene.

Despite several attempts, it was impossible to assemble the mitogenome of strain SZCZEY2172. Lowering the k-mer parameter to 75 only allowed the recovery of a short ca. 500 bp fragment with a low coverage. This fragment was used as a seed to try an assembly with NOVOPlasty 4.3.3 (Dierckxsens et al. 2017), using the mitogenome of SZCZEY2262 as reference sequence and a k-mer of 25, but this attempt also failed.

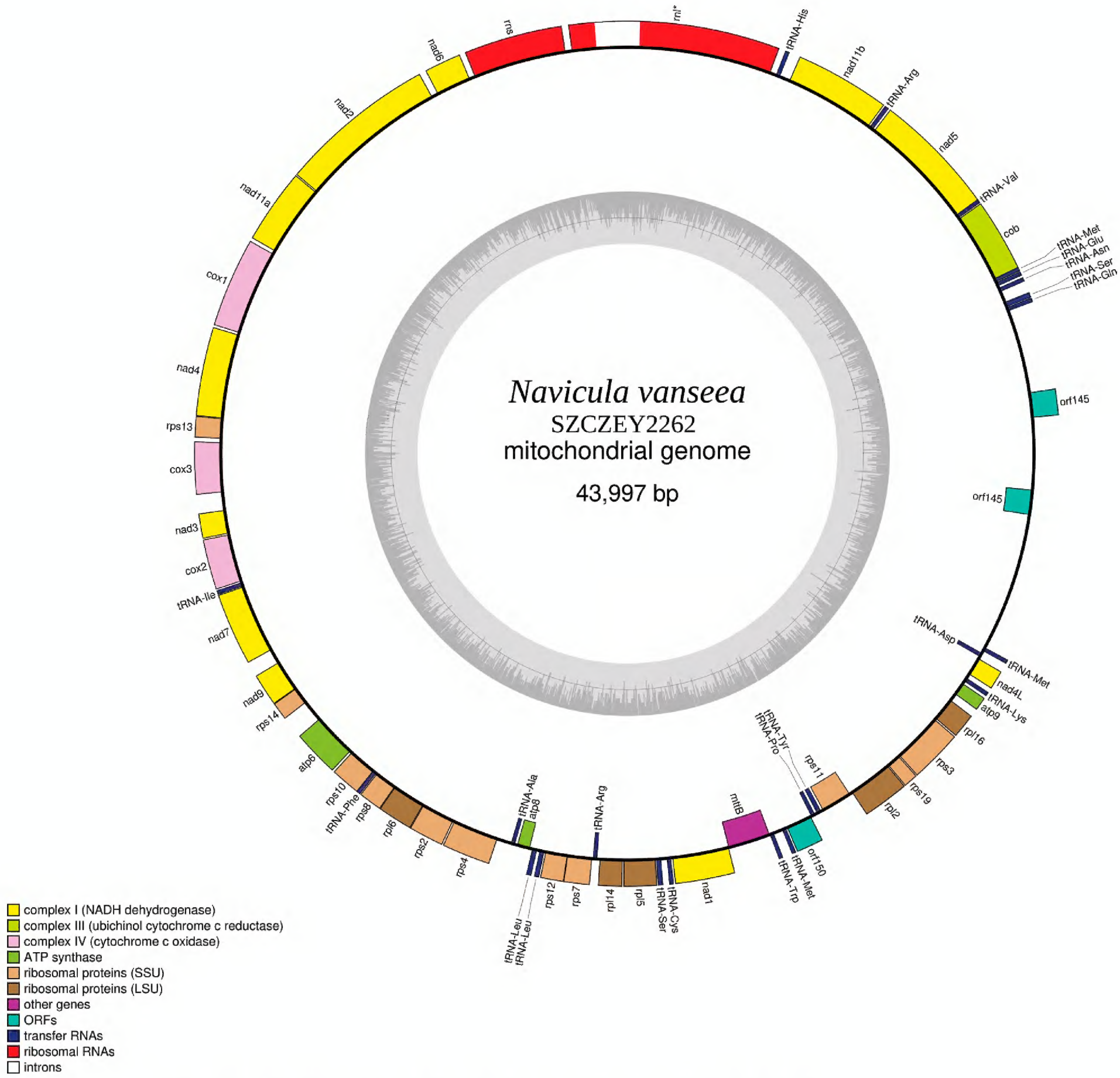

Figure 5. Map of the mitochondrial genome of *Navicula vanseea* sp. nov. SZCZEY2262.

Plastid genome

The plastome is 158,005 bp long for SZCZEY2262 (Fig. 6) and 157,990 bp long for SZCZEY2172 (Fig. 7). For SZCZEY2262 (GenBank: OR795085), the large single-copy (LSC) is 72,941 bp long and encodes 74 conserved protein-coding genes, two non-conserved ORF, two putative integrase/recombinase *xerC* genes and 17 tRNAs. The small single-copy region (SSC) is 49,714 long and encodes 51 conserved protein-coding genes, eight tRNAs, and five non-conserved ORFs of a size higher than 100 amino-acids (AA). The inverted repeat (IR) is 17,675 bp long and encodes one conserved protein-coding gene, three rRNAs, six non-conserved ORFs, four tRNAs and one putative *serC* gene and *rbcR* overlaps the inverted repeat B (IRB) and the SSC. There are two IB4 group I introns in the *rrl* gene at positions 1917 and 1931 (based on the reference sequence U00096 from *Escherichia coli* T. Escherich, 1885 str. K-12 substr. MG1655),

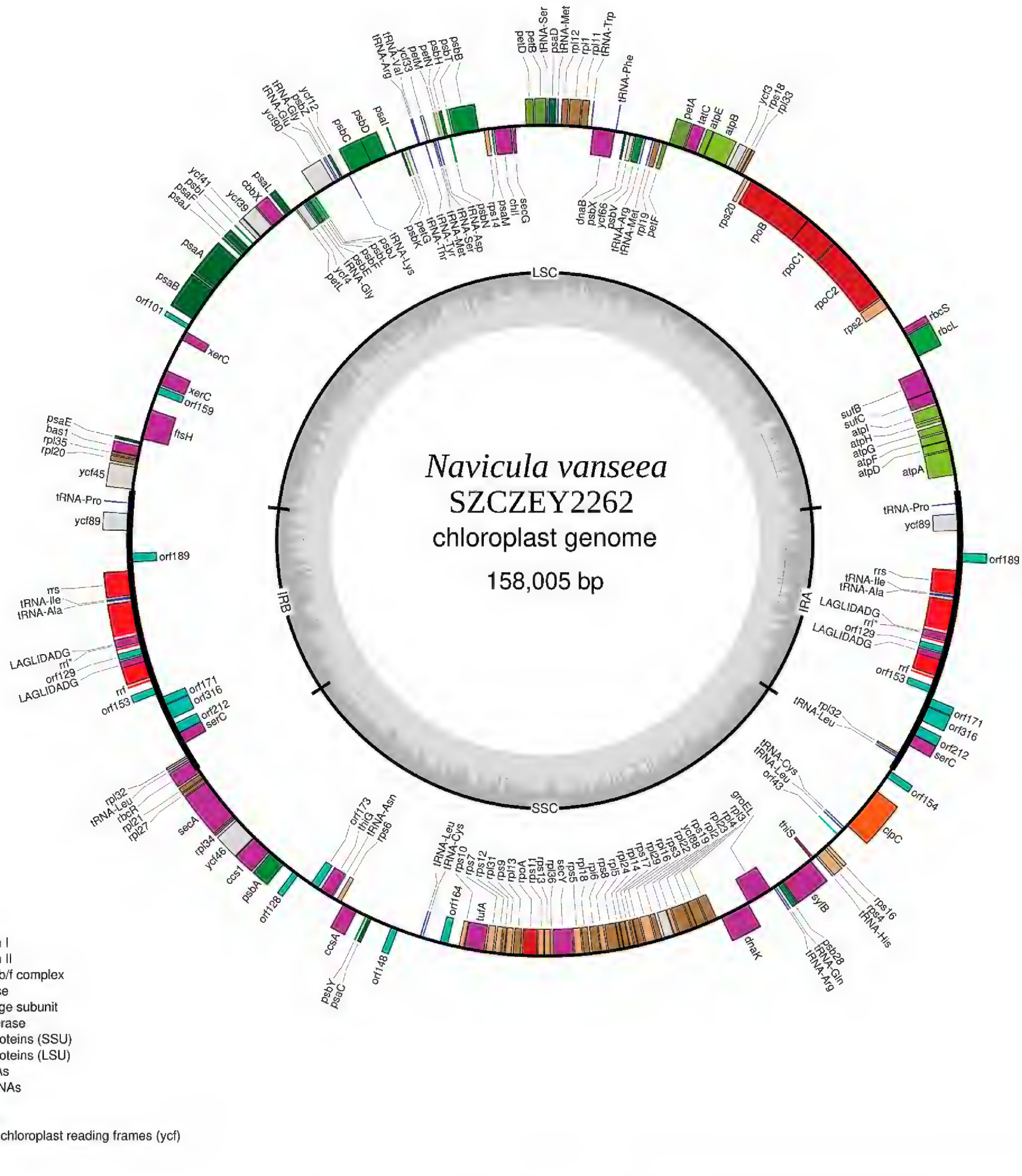

Figure 6. Map of the plastid genome of *Navicula vanseea* sp. nov. SZCZEY2262.

both containing two putative *LAGLIDADG* homing endonuclease genes. They will be referred to as L1917 and L1931.

For SZCZEY2172 (GenBank: OR795086), the LSC is 72,913 bp long and has an identical gene content compared to SZCZEY2262, the SSC is 49727 bp long and encodes 51 conserved protein-coding genes, eight tRNAs and six non-conserved ORFs of more than 100 AA. The IR is 17,675 bp long and has an identical gene and intron content to SZCZEY2262, with the same overlap of *rbcR* between the IRB and the SSC.

Both genomes contain a 43 AA ORF in their SSC that cannot be extended because of the presence of stop-codons. This ORF shows similarities to the hypothetical chloroplast RF35 encoded by *ycf35*, which is missing between both

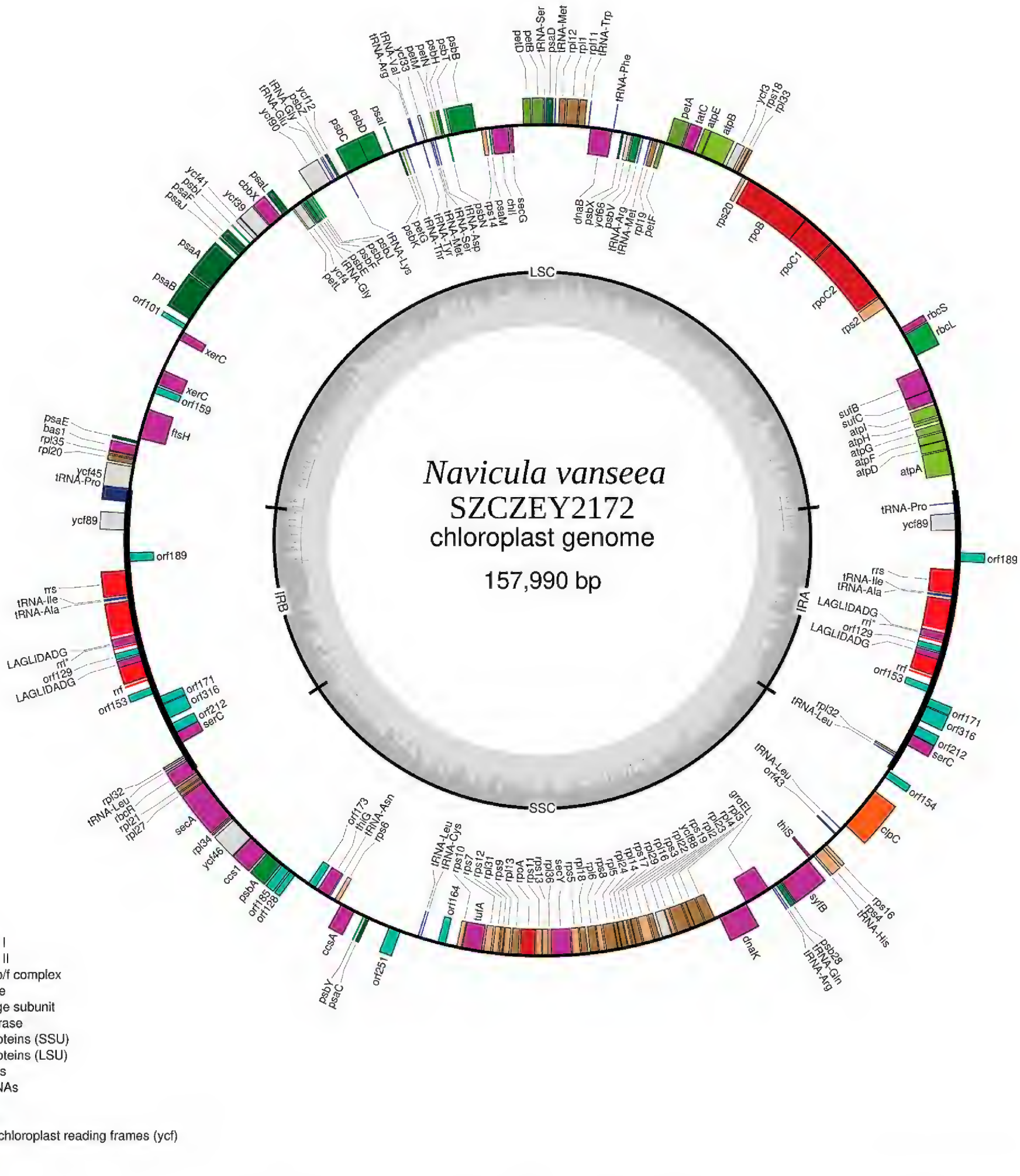

Figure 7. Map of the plastid genome of *Navicula vanseea* sp. nov. SZCZEY2172.

strains. The position of this ORF also corresponds to the position of *ycf35* in *Navicula veneta*, between *clpC* and *rps13* (Gastineau et al. 2021a).

It is worth noting that, in addition to the differences in length and content in the non-conserved ORFs, there is a slight degree of extra polymorphism in the two strains, the extent depending on the part of the genome considered. There were only two SNPs in the inverted repeat (one in the spacer between *ycf45* and *tRNA-Pro* and the other inside *rbcR*). On the other hand, a gene such as *psbC* displayed three SNPs, two of them silent, but one leading to a phenylalanine–leucine substitution. The two *xerC* genes, although present in both strains, differed in length.

Multigene phylogeny

The subtree containing Naviculaceae (Fig. 8) has been extracted from the complete multigene tree (downloadable as explained in the data availability statement). *Navicula vanseea* sp. nov. appears in a well-supported (99%) clade that also contains the freshwater diatom *Navicula cryptocephala* UTEX FD109 (sometimes indexed as *Navicula cryptocephala var. veneta* UTEX FD109), a specimen isolated by the late David B. Czarnecki from North Dakota, USA (Theriot et al. 2010) and the marine *Navicula* sp. KSA2015 41, which originates from the vicinity of Rabigh on the Red Sea, Saudi Arabia (Sabir et al. 2018). The clade also contains *Pseudogomphonema* sp. and two species of *Seminavis* spp. The phylogeny unambiguously separates *N. vansee* sp. nov. from the two morphologically similar species *N. cincta* and *N. microdigitoradiata. Navicula cincta* appears in a different, strongly-supported clade (97%) that contains *N. capitatoradiata*, *Navicula tsukamotoi* (Sterrenburg & Hinz) Yuhang Li & Kuidong Xi 2017, several unnamed species of *Navicula* spp. and *Rhoikoneis pagoensis* Lobban, 2015. *Navicula microdigitoradiata* is also easily distinguished and appears as sister to *Navicula hippodontofallax* Witkowski & Chulian Li 2016.

Phylogeny of the putative homing endonuclease LAGLIDADG proteins

Once rooted with sequence ABR25263, the phylogenetic tree of LAGLIDADG proteins (Fig. 9) distinguished the two groups. The tree associates sequences from *N. vanseea* SZCZEY2172 and SZCZEY2262 with those in other species that are of the same type and occupy the same positions. For example, the L1931 LAGLIDADG sequences of the two *N. vanseea* clones were found to be sister to a L1931 LAGLIDADG in the plastid genome of the diatom *Schizostauron trachyderma* (F. Meister) Górecka, Riaux-Gobin & Witkowski, 2021 (Górecka et al. 2021b) and then to the green algae *Pterosperma cristatum* Schiller, 1925 (Prasinophyceae) and *Pedinomonas tuberculata* (Vischer) Gams, 1947 (Pedinophyceae), a synonym of *Chlorochytridion tuberculatum* Vischer 1945 (both from plastid genomes). In contrast to the topology of the L1931 clade, in the L1917 tree, the *N. vanseea* LAGLIDADG sequences appeared at the base of the clade with maximum support. In this clade, sequences from the plastomes of various Viridiplantae form a strong clade, separated from *N. vanseea* by two Prokaryota, namely the heterotrophic bacteria *Pseudothermotoga thermarum* (Windberger et al. 1992) Bhandari and Gupta 2014 and the cyanobacterium *Synechococcus* sp. C9.

Discussion

Morphological comparison with similar taxa

Navicula vanseea sp. nov. has elliptic valves that taper towards cuneately rounded apices in smaller specimens and linear-elliptic-lanceolate with narrowly rounded, protracted endings in larger specimens (Table 1). *Navicula cincta*, *Navicula dealpina* Lange-Bert., *N. microdigitoradiata*, *N. meulemansii* and *Navicula veronensis* Lange-Bertalot and Cantonati are the most similar taxa. Regarding valve outline, *Navicula dealpina* is very similar, being linear-lanceolate to almost elliptic.

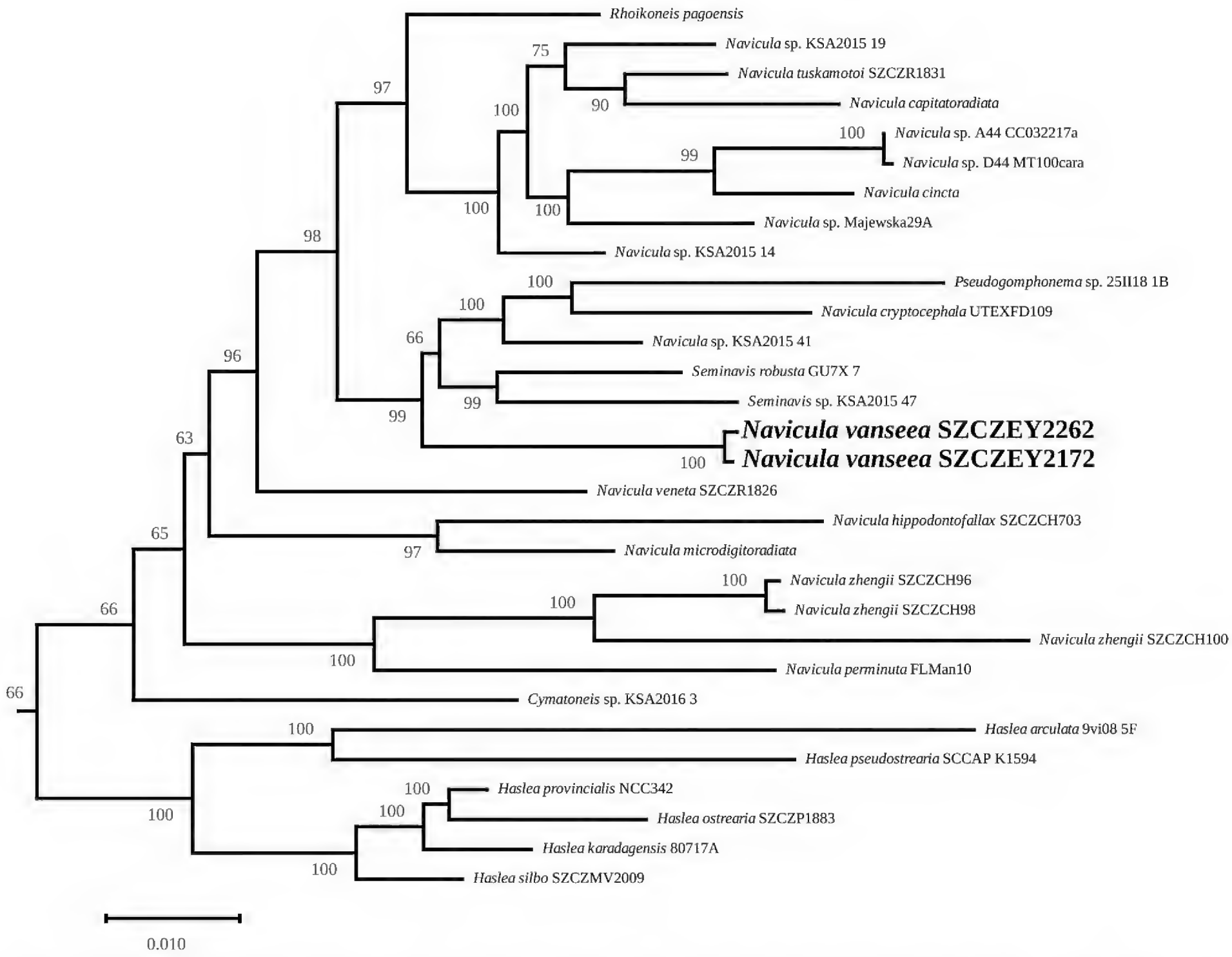

Figure 8. Maximum Likelihood phylogenetic tree obtained from concatenated alignments of *psbC*, *rbcL* and *18S*.

However, the apices of *N. dealpina* are much narrower compared to the bluntly rounded apices of *N. vanseea*. Moreover, *N. dealpina* is larger (26–86 µm length, 8–12 µm width) and has a lower stria density (8–10 striae in 10 µm) than *N. vanseea* (11–28 µm length, 5–6 µm width and 12–13 striae in 10 µm) and it has a transversely rectangular central area (not elliptic and small in *N. vanseea*). Regarding dimensions, *N. veronensis* and *N. meulemansii* are quite similar. *Navicula veronensis* has similarly linear-lanceolate valve outline, but *N. vanseea* is more elliptic. Moreover, *N. veronensis* has relatively larger and more visible central area with gradually wedge-shaped, finally obtusely rounded apices, while *N. vanseea* has an indistinctive central area with narrowly rounded apices. In SEM, *N. veronensis* has a thicker sternum structure than *N. vanseea*, while *N. vanseea* has more strongly-hooked distal raphe endings. Similarly, *N. meulemansii* have thicker sternum structure in SEM. Additionally, *N. meulemansii* has elliptic-lanceolate valve outline with cuneate apices and higher striae density than *N. vanseea* (15–21 in 10 µm vs. 12–13 in 10 µm, respectively). Amongst the similar taxa, *Navicula microdigitoradiata* and *N. cincta* have slightly larger sizes (up to 40–45 µm length and 7–8 µm width) and lower striae density (max. 10 in 10 µm for *N. cincta* and 11 in 10 µm for *N. microdigitoradiata*) than *N. vanseea* (max. 28 µm length, 6 µm width and 13 striae in 10 µm). *N. meulemansii* has an almost similar length with narrower specimens (3–5 µm width) and higher striae density (15–21 stri-

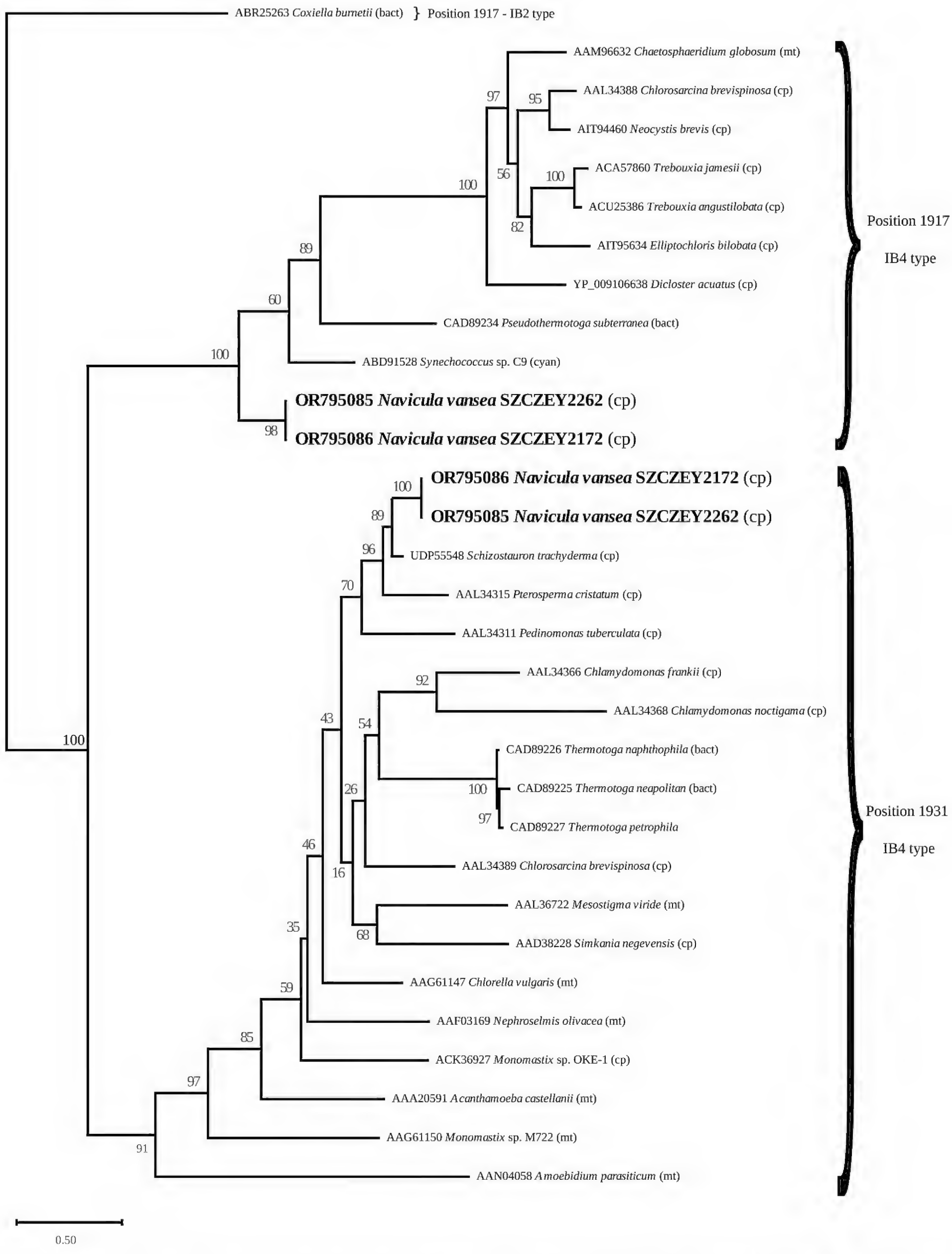

Figure 9. Maximum Likelihood phylogenetic tree inferred from the alignment of the putative LAGLIDADG endonuclease proteins found in the group I introns of *Navicula vanseea* sp. nov. and other taxa. The type of genome is indicated between brackets: cp – plastome, mt – mitogenome, bact – bacteria, cyan – cyanobacteria.

ae in 10 µm). The valve outline is relatively more elliptic with cuneately rounded apices in *N. meulemansii*, while *N. vanseea* has a more linear-lanceolate outline with narrowly-rounded apices (Table 1). It is worth mentioning that *N. cincta*, *N. veronensis* and *N. dealpina* are also all species suspected to prefer alkaline water (Cantonati et al. 2016), so it is especially important to clarify the distinctions between them and *N. vanseea*.

In relation to the three taxa mentioned in the introduction as having been found in Van Lake by Legler and Krasske (1940) – *N. cryptocephala*, *N. capitatoradiata* and *N. veneta* – *N. vansee* is easily distinguished in LM by the shape of its apices, which are rounded, while all the others have narrow, protracted apices. Molecular phylogeny, despite the limitations in the sampling of taxa, also concurs to discriminate *N. vanseea* sp. nov. from these three species (Fig. 8).

Genetic polymorphisms and genome evolution

The organellar genomes, especially the plastomes, show some interesting features. For example, introns are not considered to be conserved genetic elements and are known to vary amongst isolates of a single species (e.g. Gastineau et al. (2021b)), but they were fully conserved between our two isolates, whereas protein-coding genes displayed non-silent polymorphisms. In addition, all three markers commonly used for phylogeny reconstruction in diatoms (18S, *rbcL*

Table 1. Comparison of *Navicula vanseea* sp. nov. and similar species.

	***Navicula vanseea* sp. nov.**	***Navicula meulemansii***	***Navicula microdigitoradiata***	***Navicula cincta***	***Navicula cariocincta***	***Navicula veronensis***	***Navicula dealpina***
Valve length (µm)	11.5–28.5	12–30	15–40	14–45	30–50	19–40	26–86
Valve width (µm)	5–6	3.5–5.5	5–7	5.5–8	5.5–7	4–7	8–12
Stria density (in 10 µm)	12–13	15–21	10–11	8–10	10–12	11–13	8–10
Striae around central area	irregularly shortened	irregularly shortened	single shortened on either side	irregularly shortened	irregularly shortened	2–4 striae distinctly shortened on either side	irregularly shortened
Lineolae density (in 10 µm)	ca. 50	ca. 44–51	–	ca. 40	ca. 30	ca. 50	ca. 26 (LM)
Valve shape	smaller elliptic, larger linear-elliptic-lanceolate	elliptic-lanceolate	elliptic-lanceolate to linear-lanceolate	elliptic to lanceolate to linear-elliptic-lanceolate	linear- elliptic -lanceolate	linear-lanceolate	linear-lanceolate to almost elliptical
Central area	elliptic and small	very small and asymmetric	elliptic and very small	small	large, transversely rectangular to elliptical	broadly rectangular or transapically elliptic	almost symmetrical, transversally rectangular
Valve apices	narrowly rounded	cuneately rounded	obtusely rounded	obtusely rounded	narrowed to a wedge, obtusely-rounded, never protracted	gradually wedge-shaped, finally obtusely rounded	obtusely wedge-shaped
Raphe	filiform	weakly lateral; central pores very close together	weakly lateral, central pores very close	filiform	strongly radiate	filiform	weakly to strongly lateral, lying outside the median area towards the central pores
References	this study	Mertens et al. (2014)	Lange-Bertalot (2001)	Lange-Bertalot (2001)	Tsarenko et al. (2000)	Cantonati et al. (2016)	Lange-Bertalot (2001)

and *psbC*) exhibited single nucleotide polymorphisms in *N. vanseea*; this was especially surprising for the nuclear 18S gene, which generally exhibits very few differences between closely-related species (Evans et al. 2007). The polymorphism in *18S* appeared to be in the variable V2 region, while the V4 or V9 regions, which are often used in metabarcoding studies, were found fully conserved.

In the plastome of *N. vanseea*, the *ycf35* gene has seemingly been turned to a pseudogene, which would be the first time to our knowledge that this has been observed in diatoms, although *ycf35* pseudogenes have been observed in Rhodophyta (Costa et al. 2016). This gene seems to be lost altogether amongst other taxa, such as *Rhizosolenia imbricata* Brightwell, 1858 (Sabir et al. 2014) or *Proboscia* sp. and *Licmophora* sp. (Yu et al. 2018). It is not clear if the gene has been completely lost in these species or if it has been transferred to the nuclear genome, which is known to have happened with the plastid *petF* gene in *Thalassiosira oceanica* (Lommer et al. 2010). In *N. vanseea*, the *ycf35* gene is likely no longer functional. The size of the Ycf35 protein amongst diatoms is ca. 1130 AA long. Its origin can be traced back to Cyanobacteria. Its function is unknown, but it has been suggested, based on experiments conducted on *Synechocystis*, to participate in CO_2 capture (Jiang et al. 2015). The results obtained from *N. vansee* sp. nov. suggest that it is not necessary to its metabolism and survival, unless *ycf35* is also already present in the nucleus.

Our study also illustrates the added value that next generation sequencing provides when describing new species, in three ways. First, it is a convenient way to gather data for multigene phylogenies, whatever the species considered. Second, in the current case with *N. vanseea*, it made it possible to find SNPs in supposedly conserved genes of two sympatric strains of the same species. This needs to be taken into account in interpreting phylogenetic and metabarcoding analyses. Third, serendipitous discoveries can occur that increase our knowledge of the organellar genomes of diatoms and other stramenopiles, such as the loss of a functional *ycf35* gene here or the first documented L1917 intron found in a stramenopile.

Acknowledgements

This work started under the supervision of our beloved colleague and mentor, the late Prof. Dr. Hab. Andrzej Witkowski, who passed away on 17 September 2023. His guidance is already deeply missed.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

This work was co-financed by the Minister of Science under the “Regional Excellence Initiative” Program (2024–2027). Claude Lemieux and Monique Turmel were supported by grant RGPIN-2017-04506 from the Natural Sciences and Engineering Research Council

of Canada (NSERC). We also acknowledge support from the CERCA Programme/Generalitat de Catalunya. The Royal Botanic Garden Edinburgh is supported by the Scottish Government's Rural and Environment Science and Analytical Services Division.

Author contributions

Conceptualisation: EY. Data curation: EY. Funding acquisition: AW. Investigation: EY, RG, CNS, EG, CL, MT, NE. Methodology: AW, RT, DGM. Project administration: CNS, AW. Supervision: AW, DGM, RT, CNS, RG. Visualisation: EY, NE, CNS, EG. Writing - original draft: EY. Writing - review and editing: DGM, RT, CL, MT, CNS, RG.

Author ORCIDs

Elif Yılmaz https://orcid.org/0000-0001-7939-1814
David G. Mann https://orcid.org/0000-0003-0522-6802
Romain Gastineau https://orcid.org/0000-0001-8661-5118
Rosa Trobajo https://orcid.org/0000-0001-9498-3797
Cüneyt Nadir Solak https://orcid.org/0000-0003-2334-4271
Ewa Górecka https://orcid.org/0000-0003-0590-7480
Monique Turmel https://orcid.org/0000-0001-7060-035X
Claude Lemieux https://orcid.org/0000-0001-9580-8042
Nesil Ertorun https://orcid.org/0000-0001-6224-7314
Andrzej Witkowski https://orcid.org/0000-0003-1714-218X

Data availability

All sequences have been deposited and are already available on GenBank with accession numbers OR797293, OR797294, OR795084, OR795085, OR795086. The alignment, partition file and complete tree in Newick format can be obtained from Zenodo following this link: https://doi.org/10.5281/zenodo.10518939.

References

Adrienne Mertens AW, Lange-Bertalot H, Ribeiro L, Rhiel E (2014) *Navicula meulemansii* sp. nov., (Bacillariophyaceae) from brackish waters in Europe and the USA. Nova Hedwigia 98(1–2): 201–212. https://doi.org/10.1127/0029-5035/2013/0152

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455–477. https://doi.org/10.1089/cmb.2012.0021

Bory de Saint-Vincent JBGM (1822) Bacillariées. Dictionnaire classique d'histoire naturelle, 2. Rey et Gravier, Libraries-Editeurs, Paris, 127–129.

Cantonati M, Angeli N, Spitale D, Lange-Bertalot H (2016) A new *Navicula* (Bacillariophyta) species from low-elevation carbonate springs affected by anthropogenic disturbance. Fottea 16(2): 255–265. https://doi.org/10.5507/fot.2016.013

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics (Oxford, England) 25(15): 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Costa JF, Lin SM, Macaya EC, Fernández-García C, Verbruggen H (2016) Chloroplast genomes as a tool to resolve red algal phylogenies: A case study in the Nemaliales. BMC Evolutionary Biology 16(1): 205. https://doi.org/10.1186/s12862-016-0772-3

Cox EJ (1979) Taxonomic studies on the diatom genus *Navicula* Bory. The typification of the genus. Bacillaria 2: 137–153.

Cox EJ (1995) Studies on the diatom genus *Navicula* Bory. VII. The identity and typification of *Navicula gregaria* Donkin, *N. cryptocephala* Kütz. and related taxa. Diatom Research 10(1): 91–111. https://doi.org/10.1080/0269249X.1995.9705330

Dąbek P, Ashworth MP, Witkowski A, Li C, Bornman TG, Gonçalves V, Park J, Khim JS (2017) Towards a multigene phylogeny of the Cymatosiraceae (Bacillariophyta, Mediophyceae) I: Novel taxa within the subfamily cymatosiroideae based on molecular and morphological data. Journal of Phycology 53(2): 342–360. https://doi.org/10.1111/jpy.12501

Dąbek P, Gastineau R, Bornman TG, Lemieux C, Turmel M, Hallegraeff G, Mouget J-L, Witkowski A (2022) The blue diatom *Haslea ostrearia* from the Indian Ocean coast of South Africa, with comparative analysis of *Haslea* organellar genomes. Frontiers in Marine Science 9: 950716. https://doi.org/10.3389/fmars.2022.950716

Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. Molecular Biology and Evolution 37(1): 291–294. https://doi.org/10.1093/molbev/msz189

Dierckxsens N, Mardulyn P, Smits G (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Research 45(4): e18. https://doi.org/10.1093/nar/gkw955

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus (San Francisco, Calif.) 12: 13–15.

Ersoy Omeroglu E, Sudagidan M, Yurt MNZ, Tasbasi BB, Acar EE, Ozalp VC (2021) Microbial community of soda Lake Van as obtained from direct and enriched water, sediment and fish samples. Scientific Reports 11(1): 18364. https://doi.org/10.1038/s41598-021-97980-3

Evans KM, Wortley AH, Mann DG (2007) An assessment of potential diatom "barcode" genes (*cox1*, *rbcL*, *18S* and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158(3): 349–364. https://doi.org/10.1016/j.protis.2007.04.001

Gastineau R, Hamedi C, Baba Hamed MB, Abi-Ayad S-ME-A, Bąk M, Lemieux C, Turmel M, Dobosz S, Wróbel RJ, Kierzek A, Lange-Bertalot H, Witkowski A (2021a) Morphological and molecular identification reveals that waters from an isolated oasis in Tamanrasset (extreme South of Algerian Sahara) are colonized by opportunistic and pollution-tolerant diatom species. Ecological Indicators 121(6): 107104. https://doi.org/10.1016/j.ecolind.2020.107104

Gastineau R, Hansen G, Poulin M, Lemieux C, Turmel M, Bardeau JF, Leignel V, Hardivillier Y, Morançais M, Fleurence J, Gaudin P, Méléder V, Cox EJ, Davidovich NA, Davidovich OI, Witkowski A, Kaczmarska I, Ehrman JM, Soler Onís E, Quintana AM, Mucko M, Mordret S, Sarno D, Jacquette B, Falaise C, Séveno J, Lindquist NL, Kemp Jr PS, Eker-Develi E, Konucu M, Mouget JL (2021b) *Haslea silbo*, A Novel Cosmopolitan Species of Blue Diatoms. Biology (Basel) 10(4): 328. https://doi.org/10.3390/biology10040328

Gastineau R, Li C, Ashworth MP, Witkowski A, Turmel M, Górecka E, Frankovich TA, Wachnicka A, Lobban CS, Theriot EC, Otis C, Dąbek P, Binczewska A, Lemieux C (2022) A New Insight Into *Amicula*, a Genus of Tiny Marine Benthic Diatoms With the Description of Two New Tropical Species and the Largest Mitogenome Known for a Stramenopile. Frontiers in Marine Science 9: 941506. https://doi.org/10.3389/fmars.2022.941506

Glombitza C, Stockhecke M, Schubert CJ, Vette A, Kallmeyer J (2013) Sulfate reduction controlled by organic matter availability in deep sediment cores from the saline, alkaline Lake Van (Eastern Anatolia, Turkey). Frontiers in Microbiology 4: 209. https://doi.org/10.3389/fmicb.2013.00209

Gordon D, Green P (2013) Consed: A graphical editor for next-generation sequencing. Bioinformatics (Oxford, England) 29(22): 2936–2937. https://doi.org/10.1093/bioinformatics/btt515

Górecka E, Ashworth MP, Davidovich N, Davidovich O, Dąbek P, Sabir JSM, Witkowski A (2021a) Multigene phylogenetic data place monoraphid diatoms *Schizostauron* and *Astartiella* along with other fistula-bearing genera in the Stauroneidaceae. Journal of Phycology 57(5): 1472–1491. https://doi.org/10.1111/jpy.13192

Górecka E, Gastineau R, Davidovich NA, Davidovich OI, Ashworth MP, Sabir JSM, Lemieux C, Turmel M, Witkowski A (2021b) Mitochondrial and Plastid Genomes of the Monoraphid Diatom *Schizostauron trachyderma*. International Journal of Molecular Sciences 22(20): 11139. https://doi.org/10.3390/ijms222011139

Guillard RRL (1975) Culture of phytoplankton for feeding marine invertebrates. In: Smith WL, Chanley MH (Eds) Culture of Marine Invertebrate Animals. Plenum Press, New York, 26–60. https://doi.org/10.1007/978-1-4615-8714-9_3

Haugen P, Bhattacharya D (2004) The spread of LAGLIDADG homing endonuclease genes in rDNA. Nucleic Acids Research 32(6): 2049–2057. https://doi.org/10.1093/nar/gkh520

Haugen P, Bhattacharya D, Palmer JD, Turner S, Lewis LA, Pryer KM (2007) Cyanobacterial ribosomal RNA genes with multiple, endonuclease-encoding group I introns. BMC Evolutionary Biology 7(1): 159. https://doi.org/10.1186/1471-2148-7-159

Jiang HB, Song WY, Cheng HM, Qiu BS (2015) The hypothetical protein Ycf46 is involved in regulation of CO_2 utilization in the cyanobacterium *Synechocystis* sp. PCC 6803. Planta 241(1): 145–155. https://doi.org/10.1007/s00425-014-2169-0

Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, Griffiths-Jones S, Toffano-Nioche C, Gautheret D, Weinberg Z, Rivas E, Eddy SR, Finn RD, Bateman A, Petrov AI (2021) Rfam 14: Expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Research 49(D1): D192–D200. https://doi.org/10.1093/nar/gkaa1047

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Lange-Bertalot H (2001) *Navicula sensu stricto* 10 genera separated from *Navicula sensu lato*, *Frustulia*. A.R.G. Gantner Verlag K.G, Ruggell.

Legler F, Krasske G (1940) Diatomeen aus dem Vansee (Armenien). Beihefte zur Botanischen Zentralblatt 60, Abteilung B, 335–345.

Lemieux C, Otis C, Turmel M (2014) Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. BMC Evolutionary Biology 14(1): 211. https://doi.org/10.1186/s12862-014-0211-2

Li C, Ashworth MP, Mackiewicz P, Dąbek P, Witkowski J, Górecka E, Krzywda M, Witkowski A (2020) Morphology, phylogeny, and molecular dating in Plagiogrammaceae family focused on *Plagiogramma-Dimeregramma* complex (Urneidophycidae, Bacillariophyceae). Molecular Phylogenetics and Evolution 148: 106808. https://doi.org/10.1016/j.ympev.2020.106808

Lohse M, Drechsel O, Kahlau S, Bock R (2013) OrganellarGenomeDRAW – A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing

expression data sets. Nucleic Acids Research 41(W1): W575–W581. https://doi.org/10.1093/nar/gkt289

Lommer M, Roy AS, Schilhabel M, Schreiber S, Rosenstiel P, LaRoche J (2010) Recent transfer of an iron-regulated gene from the plastid to the nuclear genome in an oceanic diatom adapted to chronic iron limitation. BMC Genomics 11(1): 718. https://doi.org/10.1186/1471-2164-11-718

Lucas P, Otis C, Mercier JP, Turmel M, Lemieux C (2001) Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. Nucleic Acids Research 29(4): 960–969. https://doi.org/10.1093/nar/29.4.960

Millie DF, Wee JL (1981) Observations on the tube-dwelling diatom *Navicula tripunctata* var. *schizonemoides* (VH) Patr. using light and electron microscopy. Proceedings of the Iowa Academy of Science 88(2): 74–78. https://scholarworks.uni.edu/pias/vol88/iss2/8

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37(5): 1530–1534. https://doi.org/10.1093/molbev/msaa015

Patrick R (1959) New species and nomenclatural changes in the genus *Navicula* (Bacillariophyceae). Proceedings of the Academy of Natural Sciences of Philadelphia 111: 91–108. https://www.jstor.org/stable/4064508

Pogoda CS, Keepers KG, Hamsher SE, Stepanek JG, Kane NC, Kociolek JP (2019) Comparative analysis of the mitochondrial genomes of six newly sequenced diatoms reveals group II introns in the barcoding region of *cox1*. Mitochondrial DNA, Part A, DNA Mapping, Sequencing, and Analysis 30(1): 43–51. https://doi.org/10.1080/24701394.2018.1450397

Raghavan R, Miller SR, Hicks LD, Minnick MF (2007) The unusual 23S rRNA gene of *Coxiella burnetii*: Two self-splicing group I introns flank a 34-base-pair exon, and one element lacks the canonical omegaG. Journal of Bacteriology 189(18): 6572–6579. https://doi.org/10.1128/JB.00812-07

Reimer A, Landmann G, Kempe S (2009) Lake Van, eastern Anatolia, hydrochemistry and history. Aquatic Geochemistry 15(1): 195–222. https://doi.org/10.1007/s10498-008-9049-9

Round FE, Crawford RM, Mann DG (1990) The Diatoms. Biology & morphology of the genera. Cambridge Univ. Press, Cambridge, 747 pp.

Sabir JS, Yu M, Ashworth MP, Baeshen NA, Baeshen MN, Bahieldin A, Theriot EC, Jansen RK (2014) Conserved gene order and expanded inverted repeats characterize plastid genomes of Thalassiosirales. PLOS ONE 9(9): e107854. https://doi.org/10.1371/journal.pone.0107854

Sabir JSM, Theriot EC, Manning SR, Al-Malki AL, Khiyami MA, Al-Ghamdi AK, Sabir MJ, Romanovicz DK, Hajrah NH, El Omri A, Jansen RK, Ashworth MP (2018) Phylogenetic analysis and a review of the history of the accidental phytoplankter, Phaeodactylum tricornutum Bohlin (Bacillariophyta). PLOS ONE 13(6): e0196744. https://doi.org/10.1371/journal.pone.0196744

Smith SA, Dunn CW (2008) Phyutility: A phyloinformatics tool for trees, alignments and molecular data. Bioinformatics (Oxford, England) 24(5): 715–716. https://doi.org/10.1093/bioinformatics/btm619

Solak CN, Gastineau R, Lemieux C, Turmel M, Gorecka E, Trobajo R, Rybak M, Yılmaz E, Witkowski A (2021) *Nitzschia anatoliensis* sp. nov., a cryptic diatom species from the highly alkaline Van Lake (Turkey). PeerJ 9: e12220. https://doi.org/10.7717/peerj.12220

Theriot EC, Ashworth M, Ruck E, Nakov T, Jansen RK (2010) A preliminary multigene phylogeny of the diatoms (Bacillariophyta): Challenges for future research. Plant Ecology and Evolution 143(3): 278–296. https://doi.org/10.5091/plecevo.2010.418

Tsarenko PM, Lange-Bertalot H, Stupina VV, Wasser SP (2000) Bacillariophyta. Biodiversity of cyanoprocaryotes, algae and fungi of Israel: Cyanoprocaryotes and algae of continental Israel. A.R.G. Gantner Verlag K.G, Ruggell.

Turmel M, Otis C, Lemieux C (2002) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: Insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proceedings of the National Academy of Sciences of the United States of America 99(17): 11275–11280. https://doi.org/10.1073/pnas.162203299

Yu M, Ashworth MP, Hajrah NH, Khiyami MA, Sabir MJ, Alhebshi AM, Al-Malki AL, Sabir JSM, Theriot EC, Jansen RK (2018) Chapter five – evolution of the plastid genomes in diatoms. Advances in Botanical Research 85: 129–155. https://doi.org/10.1016/bs.abr.2017.11.009